



**National University**

of Computer and Emerging Sciences

Name Nimra Saeed

Roll no 22F-3040

Section MCS-2A

Subject DS5006 Deep Learning

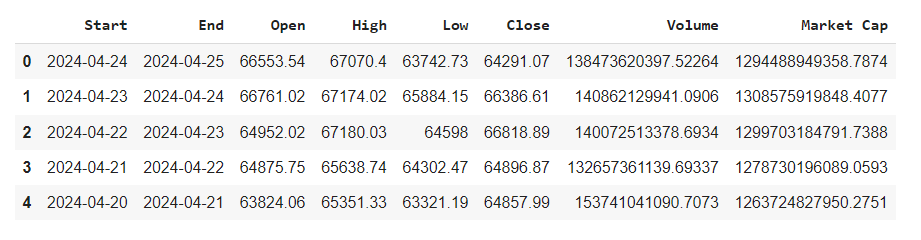
Assignment Submitted to Dr.Iqra Muhammad

**Bitcoin Historical Prices & Activity (2010-2024)**

**Dataset**

This dataset has Bitcoin's trading activity, meticulously detailing daily market statistics from July 27, 2010, to April 25, 2024. It provides a deep Bitcoin's price movements, trading volumes, and market capitalizations. The data covers over 5,021 days. Each record in the dataset consists of the following fields:

1. Start: start date of the data record.
2. End: end date of the data record.
3. Open: Bitcoin started trading at the beginning of the day.
4. High: The highest price point reached by Bitcoin during the day.
5. Low: The lowest price point reached by Bitcoin during the day.
6. Close: Bitcoin ended trading at the close of the day.
7. Volume: Bitcoin traded during the day.
8. Market Cap: Bitcoin at the end of the day.



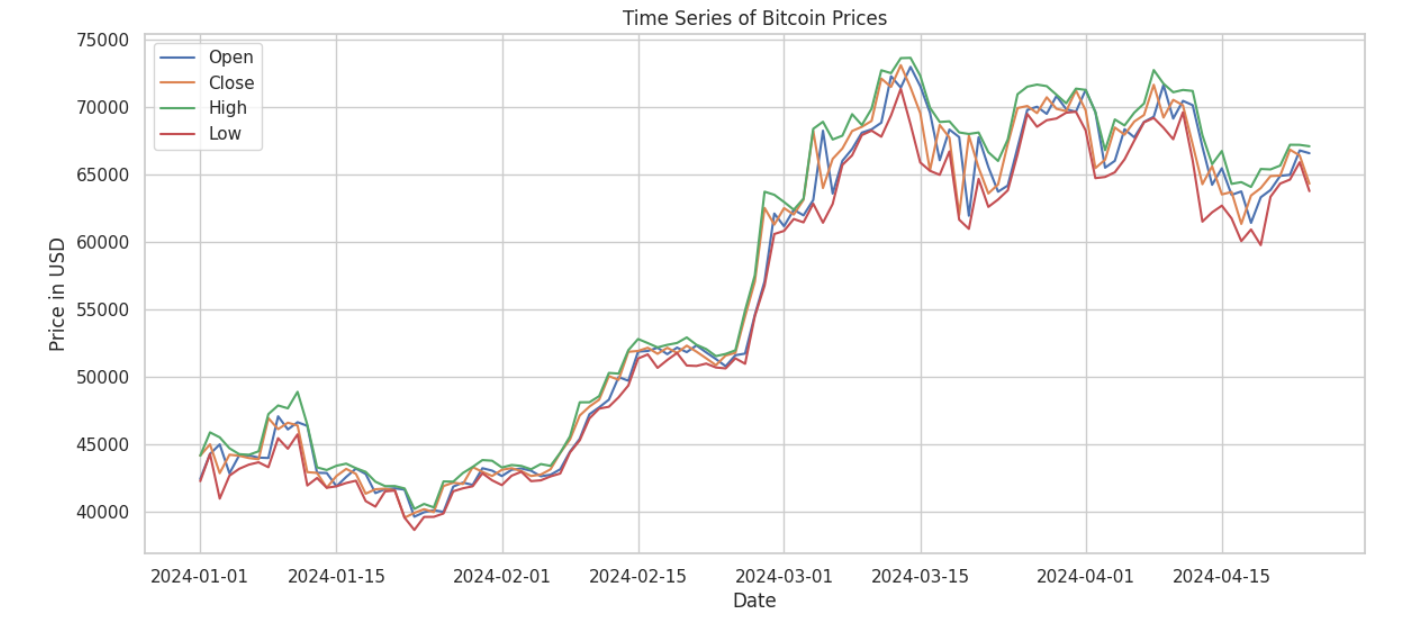
To provide a comprehensive analysis of the Bitcoin historical prices and activity dataset, explore several aspects:

* Descriptive Statistics: Overview of the price movements and trading volumes.
* Trend Analysis: How Bitcoin prices have trended over the years.
* Volatility Analysis: Examination of daily price changes.
* Volume Analysis: Relationship between volume and price changes.
* Market Cap Analysis: Trends in market capitalization over time.

Let's **focus on the year 2024**, analyze the Bitcoin dataset has 115 rows and 8 columns.

covering aspects like:

* Monthly Trends: Examining how prices have fluctuated on a monthly basis within 2024.
* Volatility: Assessing the price volatility throughout the months of 2024.
* Volume Analysis: Investigating how trading volumes have varied month by month.
* Market Cap Fluctuations: Observing changes in market capitalization.



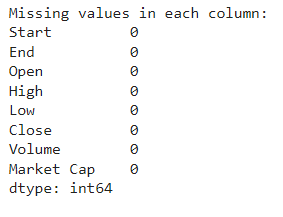
**Tools and Libraries**

* TensorFlow or PyTorch: For constructing and training deep learning models.
* Pandas and NumPy: For data manipulation and numerical calculations.
* Matplotlib/Seaborn: For generating visualizations.

**Step-by-Step Machine learning Pipeline**

**Dataset Preparation**

* Data Cleaning: Check for missing values or inconsistencies, and handle them appropriately. Ensure data types are correct for all columns.
* But in this dataset, there is no missing values.

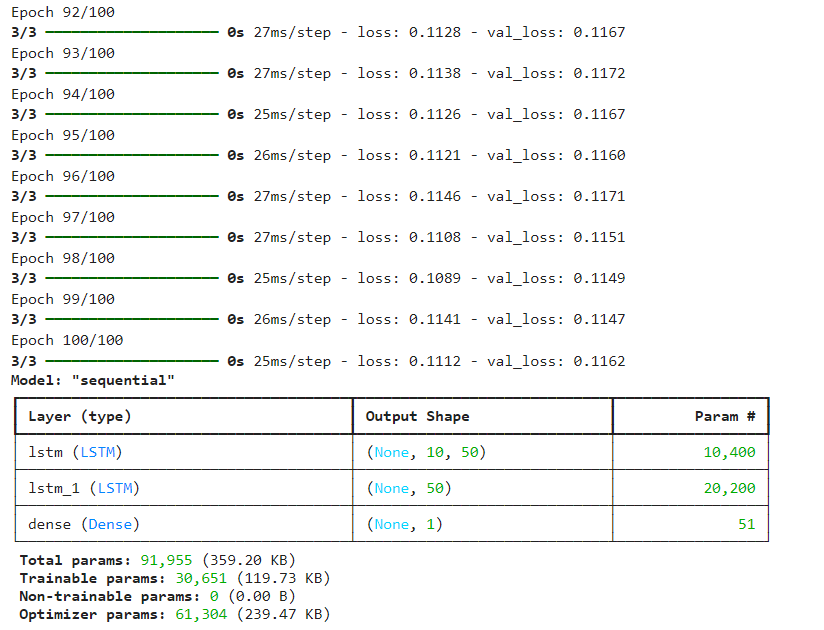


**Data Preprocessing**

* Feature Engineering: such as moving averages, percentage changes, or relative price changes.
* Normalization: Scale the features, particularly price and volume.

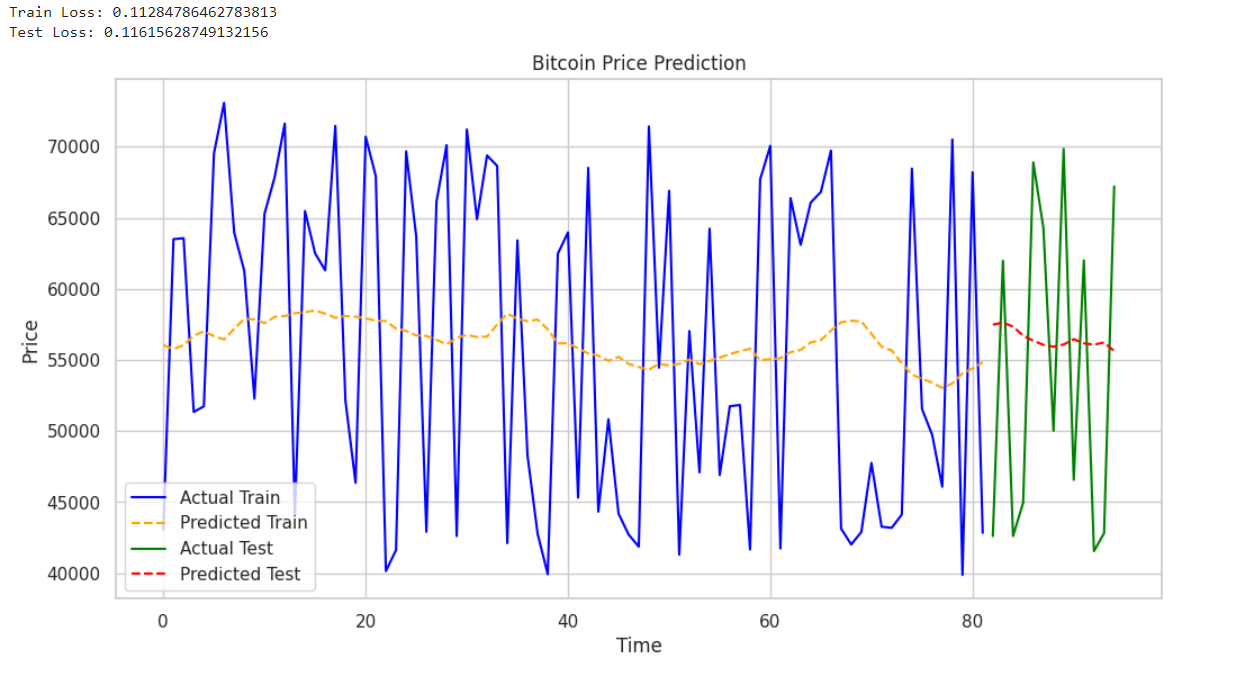
**Model Development**

* Choosing a Model: Select LSTM (Long Short-Term Memory) networks are a popular choice due to their ability to capture long-term dependencies in sequence data.
* Building the Model: Define the model architecture using TensorFlow or PyTorch. Include layers like LSTM followed by Dense layers, and consider dropout for regularization.
* Training: Train the model using hyperparameters such as learning rate, number of epochs, and batch size.



**Model Evaluation**

* Performance Metrics: Use metrics like RMSE (Root Mean Squared Error) and MAE (Mean Absolute Error) for regression tasks to evaluate the forecasting accuracy.
* Validation: Validate model on techniques like k-fold cross-validation to ensure robustness.



**Visualization**

* Plot Predictions: Visualize the predicted prices against the actual prices to compare how well your model is performing.
* Error Analysis: Plot and analyze prediction errors over time to identify any patterns or biases.

